Re: Fw: help on GOF test

Krista Christensen to: Bob Benson 01/03/2013 01:54 PM

From: Krista Christensen/DC/USEPA/US
To:
Cc: brattin@srcinc.com, David Berry/R8/USEPA/US@EPA, Leonid Kopylev/DC/USEPA/US@EPA

Hi Bob-
I should have mentioned this in the email, but I spoke with Leonid briefly and learned that he and Bill had already had some email exchanges about the utility of the HL GOF test (appropriate when, as here, we have ungrouped data on a binary outcome). Therefore I only addressed the implementation question. Please let us know if you would like to discuss further--thanks!

Krista

Bob Benson---01/03/2013 01:31:18 PM---I meant to send this message to Bill, but messed up. Can you address Bill's question about which tes

From: Bob Benson/R8/USEPA/US
To: Krista Christensen/DC/USEPA/US@EPA
Cc: brattin@srcinc.com, David Berry/R8/USEPA/US@EPA
Date: 01/03/2013 01:31 PM
Subject: Fw: help on GOF test

I meant to send this message to Bill, but messed up.

Can you address Bill's question about which test (Hosmer-Lemeshow or Chi squared) is more appropriate for GOF?

----- Forwarded by Bob Benson/R8/USEPA/US on 01/03/2013 11:26 AM -----

From: Bob Benson/R8/USEPA/US
To: Krista Christensen/DC/USEPA/US@EPA,
Date: 01/03/2013 11:25 AM
Subject: Re: help on GOF test

It is good to know that your calculation is correct. However, she didn't answer your question about which approach is better.

How should we proceed?

Krista Christensen---01/03/2013 07:42:06 AM---Hi Bill- I apologize for the late reply, just got back into the office today. I used my SAS macro o

From: Krista Christensen/DC/USEPA/US
To: "Brattin, Bill" <brattin@srcinc.com>,
Cc: Bob Benson/R8/USEPA/US@EPA, David Berry/R8/USEPA/US@EPA, Leonid Kopylev/DC/USEPA/US@EPA
Date: 01/03/2013 07:42 AM
Subject: Re: help on GOF test

Hi Bill-
I apologize for the late reply, just got back into the office today. I used my SAS macro on your test data and got the same answer you found (Hosmer-Lemeshow test statistic = 11.882523704, p-value = 0.15652). I am attaching a SAS program which demonstrates the use of SAS built-in function (available in PROC LOGISTIC) and my code to estimate the HL test statistic, and application to your test data.

[attachment "HL_GOF_Test_3Jan2013.sas" deleted by Bob Benson/R8/USEPA/US]

Note that there may be a slight difference in the built-in vs. the hard-coded version due to the method used to form the 10 groups, but where I've been able to run both for the Libby data, this hasn't affected the results. I don't think the original method proposed by H&L specified how to form the groups, just that there be '10 equally sized groups', so am not too concerned that one is more correct than the other.

Hope this helps!
Krista

"Brattin, Bill" ---12/26/2012 12:30:31 PM---Leonid and Krista I am hoping you can give me some help with regard goodness of fit testing for the

From: "Brattin, Bill" <brattin@srcinc.com>
To: Leonid Kopylev/DC/USEPA/US@EPA, Krista Christensen/DC/USEPA/US@EPA
Cc: Bob Benson/R8/USEPA/US@EPA, David Berry/R8/USEPA/US@EPA
Date: 12/26/2012 12:30 PM
Subject: help on GOF test

Leonid and Krista

I am hoping you can give me some help with regard goodness of fit testing for the exposure-response models for the Marysville cohort.

In the existing write-up of your efforts, you indicate that you used the Hosmer-Lemeshow method.
I have done a little reading on this method, and I think I know how to implement the calculations in Excel.
The attached Excel file has a test data set along with my results for the Hosmer-Lemeshow test.
Could I impose on you to check to see if you get the same answer as me?

However, in running the data set using the log-logistic model in BMDS, I have noted that BMDS does not use this approach, but rather calculates a chi-squared value for the ungrouped data, where the chi-square term is given by:

ChiSq = SUM of the following: (Obs-Pred)^2 / [pred * (1-Pred)]

In the example sheet I have implemented this method, and the results do not resemble the HL approach.

I would be very grateful if you could help me understand which method is better, and why.

****

Bill Brattin
SRC, Inc.
999 18th Street Suite 1150
Denver CO 80202

Phone: 303-357-3121
Fax: 303-292-4755
e-mail: brattin@srcinc.com
[attachment "HL Example data set.xlsx" deleted by Krista Christensen/DC/USEPA/US]